International Journal of Modern Physics A

ADVANCED TECHNOLOGIES FOR DISTRIBUTED DATABASE SERVICES HYPERINFRASTRUCTURE*

ALEXANDRE VANIACHINE

High Energy Physics Division, Argonne National Laboratory, 9700 S. Cass Ave Argonne, Illinois 60439, U.S.A.

DAVID MALON

Decision and Information Sciences Division, Argonne National Laboratory, 9700 S. Cass Ave Argonne, Illinois 60439, U.S.A.

MATTHEW VRANICAR

PIOCON Technologies, 1952 McDowell Road, Suite 104, Naperville, Illinois 60563, U.S.A.

Received (1 November 2004)

HEP collaborations are deploying grid technologies to address petabyte-scale data processing challenges. In addition to file-based event data, HEP data processing requires access to terabytes of non-event data (detector conditions, calibrations, etc.) stored in relational databases. Inadequate for non-event data delivery in these amounts, database access control technologies for grid computing are limited to encrypted message transfers. To overcome these database access limitations one must go beyond the existing grid infrastructure. A proposed hyperinfrastructure of distributed database services implements efficient secure data access methods. We introduce several technologies laying a foundation of a new hyperinfrastructure. We present efficient secure data transfer methods and secure grid query engine technologies federating heterogeneous databases. Lessons learned in a production environment of ATLAS Data Challenges are presented.

Keywords: Database; computational grid; data security; collective dynamics.

1. Introduction

ATLAS and other Large Hadron Collider (LHC) experiments are facing an unprecedented multi-petabyte event data processing task - a primary focus of the LHC Computing Grid (LCG) project. In addition to file-based event data, HEP data processing involves large amounts of valuable non-event data stored in relational databases. In contrast to the file-based data, the database-resident data flow on the grid has to be detailed further[1]. Going beyond the existing grid infrastructure we

*Work supported in part by the U.S. Department of Energy, Division of High Energy Physics, under Contract W-31-109-Eng-38.

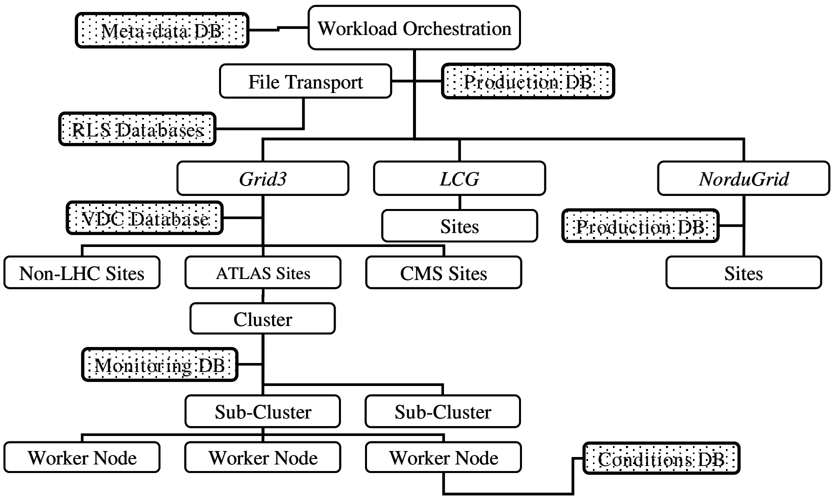

Fig. 1. A hyperinfrastructure of distributed database services crosscutting through a multi-tier hierarchy of a federation of the computational grids.

introduced several new technologies in ATLAS Data Challenges—DC1[2] and DC2[3]—processing data on a federation of three large computational grids: LCG[4], Grid3[5], and NorduGrid[6]. The new LCG Distributed Deployment of Databases project provides a testbed for further advancement of these technologies.

2. Exploring Collective Dynamics in a Federation of Grids

To effectively run production on a federation of computational grids, a single database on top does not suffice. Going beyond the existing infrastructure, a hyperinfrastructure of distributed databases crosscuts a multi-tier hierarchy of computational grids (Figure 1). The crosscutting allows separation of concerns across both the global environment of a federation of computational grids and the local environment of a physicist's laptop used for analysis.

2.1. *Cooperative Phenomena in Grid Workloads*

A data-mining of collected operations data reveals a multitude of cooperative phenomena among independent computational tasks on the grid. For example, if the computational task submitted to some cluster failed, there is a correlated probability that the next task submitted to the same cluster would fail, too. Or, if the grid user host failed, most of the tasks submitted from that host to the grid will be lost. Taking these correlations into account is not yet automated by grid middleware. On the contrary, these correlations could be further amplified, producing "black holes"[4].

In the centralized and controlled environment of traditional HEP data processing, the database services capacities required are simple to calculate. In a selforganized unpredictable environment of computational grids, the chaotic nature of grid computations increases fluctuations in database servers' workloads, making efficient capacity provisioning a challenging task. To improve estimates of the capacities required for robust operations we studied collective dynamics of the chaotic workloads observed on a federation of computational grids. For that we used database server monitoring data collected in ATLAS DC2. In DC2 Phase 1, to retrieve input parameters required by the data transformation, each computational task accessed the same database[7,8]. In these settings, when thousands of independent tasks accessed the database daily, the daily average database server load achieved 122 queries per second with a standard deviation of 0.24 of the daily average value, which is fourteen times more than purely statistical fluctuations. The workload fluctuations became comparable to the statistical ones when averaged over the period of more than 23 days. We believe that non-statistical fluctuations indicate cooperative phenomena in the collective dynamics of the computational tasks harnessing the power of more than thirteen thousand processors on a federation of three worldwide computational grids.

2.2. *High-performance Secure Databases Access*

HEP experiments are accumulating experience with secure database access technologies on the grid. Secure grid query engine technologies federating heterogeneous databases on the grid are developed for Run II experiments[9]. ATLAS utilizes a Grid Security Infrastructure data-transport channel for database services delivery to grid clusters behind closed firewalls, and uses grid certificate authorization technologies for database access control. Instead of surrounding databases with external security layers, the safety features are embedded inside of the code[3]. To avoid a brittle monolithic system, we are looking into aspect-oriented programming techniques, encapsulating grid security aspects as a separate concern. Use of the same grid security model for all data flow channels eliminates clear-text passwords and inefficient data transfer bottlenecks.

References

1. P. Watson, in *Grid Computing: Making the Global Infrastructure a Reality*, ed. F. Berman et al. (Wiley, 2003)
2. A. Vaniachine, *Nucl. Instr. and Meth.* **A502**, 446 (2003); R. Sturrock et al, European Organization for Nuclear Research, Preprint CERN-PH-EP-2004-28, 2004.
3. M. Branco et al., in *Proceedings of CHEP04*, id 189, 2004 (unpublished).
4. I. Bird et al., in *Proceedings of CHEP04*, id 186, 2004 (unpublished).
5. I. Foster et al., in *Proc. 13th Int'l Symp. on High Performance Distributed Computing* (IEEE, 2004), p. 236; M. Mambelli et al., in *Proc. of CHEP04*, id 503, 2004.
6. P. Eerola et al., in *Proc. 4th Int'l Workshop on Grid Computing* (IEEE, 2003), p. 158.
7. A. Vaniachine et al., in *Proceedings of CHEP03*, eConf C030324, MOKT006 (2003)
8. A. Vaniachine et al., in *Proceedings of CHEP04*, id 212, 2004 (unpublished).
9. http://www.piocon.com/databasegrid.html